ZooKeys 1166: 141–154 (2023)
DOI: 10.3897/zookeys.1166.104103

Research Article

Characterization of two new *Pylorgus* mitogenomes (Hemiptera, Lygaeidae, Ischnorhynchinae) and a mitochondrial phylogeny of Lygaeoidea

Cuiqing Gao[1], Wen Dong[1]

1 *Co-Innovation Center for Sustainable Forestry in Southern China, College of Forestry, Nanjing Forestry University, Nanjing, Jiangsu 210037, China*
Corresponding author: Cuiqing Gao (cqgao@njfu.edu.cn)

Abstract

Lygaeidae is a large family of Hemiptera (Heteroptera) currently separated into three subfamilies, Ischnorhynchinae, Lygaeinae, and Orsillinae. In this research, the complete mitogenomes of the iscnorhynchines *Pylorgus porrectus* Zheng, Zou & Hsiao, 1979 and *Pylorgus sordidus* Zheng, Zou & Hsiao, 1979 were sequenced, and the phylogeny of *Pylorgus* and the Lygaeidae with known complete mitogenomes were examined. The mitogenomes are 15,174 bp and 15,399 bp in size, respectively, and comprised of 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), two ribosomal RNA genes (rRNAs), and a control region (D-loop). Nucleotide composition is biased toward A and T, and the gene order is identical to that of the putative ancestral arrangement of insects. Eleven PCGs begin with a typical ATN, and the remaining two PCGs begin with TTG (*cox1* and *nad4l*). All tRNAs had a typical cloverleaf secondary structure, but some of them had individual base mismatches. The phylogenetic analyses based on the concatenated nucleotide sequences of the 13 PCGs, using Bayesian inference and maximum likelihood, support the monophyly of Lygaeidae. The results show that *P. porrectus* and *P. sordidus* clustered with nine other Lygaeidae. This study includes the first complete sequencing of the mitochondrial genomes of two *Pylorgus* species, which will provide important data for studying the phylogenetic position of Lygaeidae in Lygaeoidea and reconstructing the phylogenetic relationships within Pentatomomorpha.

Key words: Heteroptera, mitochondrial DNA, next-generation sequencing, phylogenetic analysis, *Pylorgus porrectus*, *Pylorgus sordidus*

OPEN ACCESS

Academic editor: Jader Oliveira
Received: 27 March 2023
Accepted: 24 May 2023
Published: 8 June 2023

ZooBank: https://zoobank.org/2599C45A-F635-448B-87D4-CC004535BB08

Citation: Gao C, Dong W (2023) Characterization of two new *Pylorgus* mitogenomes (Hemiptera, Lygaeidae, Ischnorhynchinae) and a mitochondrial phylogeny of Lygaeoidea. ZooKeys 1166: 141–154. https://doi.org/10.3897/zookeys.1166.104103

Copyright: © Cuiqing Gao & Wen Dong. This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Introduction

The Lygaeoidea represents the second largest superfamily within the infraorder Pentatomomorpha and includes over 4660 described species in 16 families (Henry 2017; Dellapé and Henry 2020). Most Lygaeoidea feed mainly on mature seeds (Schuh and Slater 1995); although Blissidae, Colobathristidae, Malcidae, and Piesmatidae predominantly feed on plant sap (Sweet 2000; Henry et al. 2015), Berytidae are mostly phytophagous, with a few becoming pests, although some have been shown to be predatory (Henry 2000), and Geocoridae are primarily predators but sometimes also feed on seeds and leaves of plants (Sweet 2000).

Currently, three subfamilies of Lygaeidae (sensu stricto) are recognized: Ischnorhynchinae, Lygaeinae, and Orsillinae (Dellapé and Henry 2020). The main diagnostic characters of Lygaeidae are as follows: bucculae well developed, pronotal calli with an impressed transverse groove, scutellum usually with a raised cross-shaped carina, and hamus present on wings. Abdominal spiracles on segments II to VII dorsal (Malipatil et al. 2020).

To date, the phylogeny of Lygaeidae is unresolved (Yao et al. 2012; Zhang et al. 2019), and the status of Orsillinae and Ischnorhynchinae in relation to Lygaeidae (sensu stricto) continues to be discussed. Henry (1997) proposed that Orsillinae and Ischnorhynchinae be classified as subfamilies of Lygaeidae. However, Sweet (2000) recognized them as separate families from the Lygaeidae (Orsillidae and Ischnorhynchidae). A few workers have followed Sweet in adopting the family Orsillidae (Eyles and Malipatil 2010; Malipatil 2010; Ge and Li 2019), whereas Henry et al. (2015), supported by Schuh and Weirauch (2020), disagreed with Sweet, who provided no evidence to support his hypothesis.

The complete mitochondrial genome data of nine species in Lygaeidae are included on NCBI, and only two species of Ischnorhynchinae. However, for the largest genus in this subfamily, *Pylorgus*, the mitochondrial genome data is totally unknown. Therefore, in the present study, we obtained the complete mitochondrial genomes of two *Pylorgus* species, *Pylorgus porrectus* Zheng, Zou & Hsiao, 1979 and *Pylorgus sordidus* Zheng, Zou & Hsiao, 1979, by using the next-generation sequencing technology. Furthermore, we constructed the phylogenetic trees based on the mitogenomes of 21 species of the superfamily Lygaeoidea and four outgroup species, which will provide important data for further studies on the phylogenetic position of Lygaeidae in Lygaeoidea and be also useful to reconstruct the phylogenetic relationships within Pentatomomorpha.

Materials and methods

Sample collection, DNA extraction, and mitogenome sequencing

Adults of *Pylorgus porrectus* (Fig. 1a, b) were collected from Zhongshan Botanical Garden (32°03.27'N, 118°49.85'E), Nanjing, Jiangsu Province, China, 20 April 2022, Cuiqing Gao leg. Adults of *P. sordidus* (Fig. 1c) were collected from Hongqi Management and Protection Station, Yintiaoling Nature Reserve (31°23.87'N, 109°41.32'E), Wuxi County, Chongqing, China, 1 July 2022, Suyan Cao leg. The specimens were identified based on the morphological characteristics seen under a Zeiss Stereo Discovery V8 Zoom Microscope and deposited in the Insect Collection, College of Forestry, Nanjing Forestry University.

The complete genomic DNA was extracted from an adult sample using a Rapid Animal Genomic DNA Isolation Kit (Sangon Biotech, Shanghai, China). Libraries were prepared on an Illumina MiSeq PE300 platform (Sangon Biotech, Shanghai, China). Low-quality and short reads were removed using Fastp v. 0.36 (Chen et al. 2018) to obtain clean reads and ensure rich quality of information analysis.

Mitogenome assembly, annotation, and analyses

SPAdes v. 3.15 (Bankevich et al. 2012) was used to assemble the high-quality next generation sequencing data de novo to construct contig and scaffold. After the

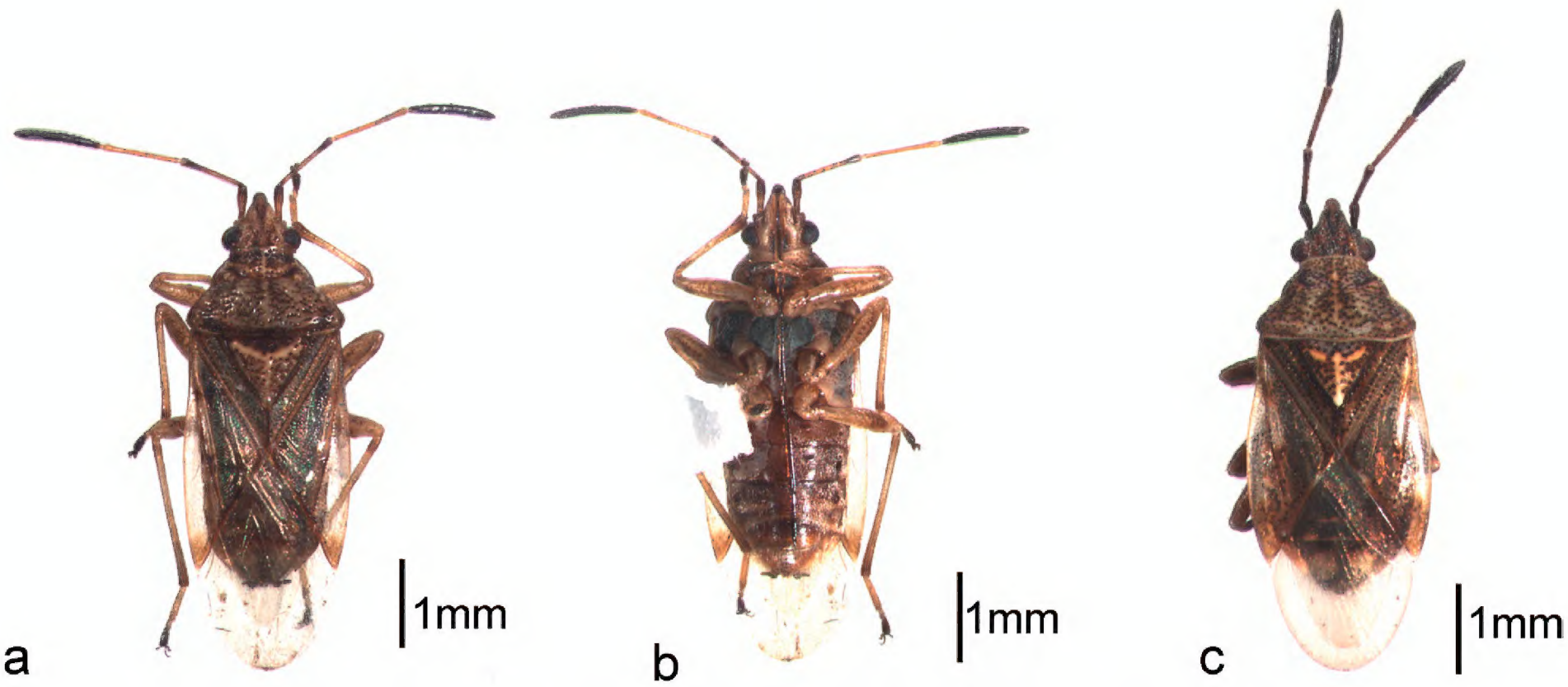

Figure 1. *Pylorgus* species sequenced **a, b** *P. porrectus*, dorsal and ventral views **c** *P. sordidus*, dorsal view.

assembly was completed, we evaluated and quality controlled the assembly results, excluding any contamination that may originate from the host genome in the subsequent analysis, and only retained the scaffolds derived from the genome of the organelle. We used BLASTn to compare the scaffolds with the NCBI library to obtain sequence similarity information, extracted the sequencing depth and coverage information of each scaffold, and manually selected possible target scaffolds after sorting out and comprehensively considering the above information. Then GapFiller v. 1.11 (Boetzer and Pirovano 2012) was adopted to supplement GAP to the contig obtained by splicing, and PrInSeS-G was adopted to carry out sequence correction to correct editing errors and insertion and deletion of small fragments in the splicing process, and finally the complete mitochondrial genome was obtained.

For mitochondrial gene annotation, we used tBLASTn and GeneWise to back-align with near-source reference databases to obtain the coding sequence (CDS) gene boundaries, and MiTFi to obtain the transfer RNA genes (tRNAs) sequence annotation. The non-coding ribosomal RNA genes (rRNAs) were identified by cm-searchrfam alignment and finally summarized into a complete annotation result.

The nucleotide composition and RSCU (relative synonymous codon usage) were calculated using PhyloSuite v. 1.2.2 (Zhang et al. 2020) and MEGA X (Kumar et al. 2018). Strand asymmetry was calculated using the formula: AT-skew = [A−T]/[A+T] and GC-skew = [G−C]/[G+C] (Perna and Kocher 1995). DnaSP v. 5 (Librado and Rozas 2009) was used to calculate the value of Ka (the nonsynonymous substitution rate), Ks (the synonymous substitution rate), and nucleotide diversity.

Phylogenetic analyses methods

The mitochondrial genome data of 25 species in Pentatomomorpha were used to reconstruct the phylogenetic relationship of Lygaeoidea, in which 21 species of Lygaeoidea were regarded as ingroup and four species was regarded as outgroup (Table 1). All sequences were standardized and extracted 13 protein-coding genes (PCGs) by PhyloSuite v. 1.2.2 (Zhang et al. 2020). The 13 PCGs of the 25 species were aligned individually using codon-based multiple alignments with MAFFT v. 7.313 software with default settings (Katoh and Standley 2013).

Table 1. Species used in this study.

Family	Subfamily	Species	Length (bp)	GenBank No.
Lygaeidae	Ischnorhynchinae	*Kleidocerys resedae* (Panzer, 1797)	14,688	KJ584365.1
	Ischnorhynchinae	*Pylorgus porrectus* Zheng, Zou & Hsiao, 1979	15,174	OP793792
	Ischnorhynchinae	*Pylorgus sordidus* Zheng, Zou & Hsiao,1979	15,399	OQ064783
	Ischnorhynchinae	*Crompus oculatus* Stål, 1874	15,332	MW619652.1
	Lygaeinae	*Arocatus melanocephalus* (Fabricius, 1798)	15,409	NC_063142.1
	Lygaeinae	*Tropidothorax cruciger* (Motschulsky, 1859)	15,781	NC_056293.1
	Lygaeinae	*Tropidothorax sinensis* (Reuter, 1888)	15,422	MW547017.1
	Orsillinae	*Nysius plebeius* Distant, 1883	17,637	MN599979.1
	Orsillinae	*Nysius cymoides* (Spinola, 1837)	16,301	MW291653.1
	Orsillinae	*Nysius fuscovittatus* Barber, 1958	14,575	NC_050167.1
	Orsillinae	*Nithecus jacobaeae* (Schilling, 1829)	14,206	MW619651.1
Berytidae	Metacanthinae	*Yemmalysus parallelus* Stusak, 1972	15,747	NC_012464.1
	Metacanthinae	*Metatropis longirostris* Hsiao, 1974	15,744	NC_037373.1
Blissidae		*Bochrus foveatus* Distant, 1879	14,738	NC_065814.1
		Capodemus sinuatus (Slater, Ashlock & Wilcox, 1969)	15,199	NC_065815.1
Geocoridae	Geocorinae	*Geocoris pallidipennis* (Costa, 1843)	14,592	NC_012424.1
	Henestarinae	*Henestaris halophilus* (Burmeister, 1835)	14,868	MW619656.1
Malcidae	Chauliopinae	*Chauliops fallax* Scott, 1874	15,739	NC_020772.1
	Malcinae	*Malcus inconspicuous* Štys, 1967	15,575	NC_012458.1
Rhyparochromidae	Rhyparochrominae	*Neolethaeus assamensis* (Distant, 1901)	15,067	NC_037375.1
	Rhyparochrominae	*Bryanellocoris orientalis* Hidaka, 1962	15,606	NC_063139.1
Pyrrhocoridae		*Dysdercus evanescens* Distant, 1902	15,635	MW619727.1
Coreidae	Hydarinae	*Hydaropsis longirostris* (Hsiao, 1963)	16,521	EU427337.1
Rhopalidae		*Aeschyntelus notatus* Hsiao, 1963	14,532	EU427333.1
Alydidae		*Riptortus pedestris* (Fabricius, 1775)	17,191	EU427344.1

Gblocks v. 0.91b software was used to remove the intergenic gaps and ambiguous sites (Talavera and Castresana 2007), and all PCGs sequences were concatenated in PhyloSuite v. 1.2.2. The best partitioning scheme and evolutionary models for constructing Bayesian inference (BI) and maximum-likelihood (ML) trees were selected by PartitionFinder2 (Lanfear et al. 2016), with a greedy algorithm, BIC criterion, and the gene and codon model.

BI phylogenies were inferred using MrBayes v. 3.2.6 (Ronquist et al. 2012) under partition model (2,000,000 generations), in which the initial 25% of sampled data were discarded as burn-in. ML phylogenies were inferred using IQ-TREE (Nguyen et al. 2015) under the Edge-linked partition model for 5000 standard bootstraps with 1000 replicates.

Results

Genome structure and composition

The assembled complete mitogenomes of *Pylorgus porrectus* and *P. sordidus* are circular DNA molecules of 15,174 bp and 15,399 bp in length, respectively,

which is within the range of the sequenced mitogenomes of Lygaeidae in GenBank (Table 1). These mitogenomes all have a similar typical insect mitogenome structure, closed-circular and double-stranded DNA, containing 13 PCGs, 22 tRNAs, two rRNAs, and a control region (D-loop) (Fig. 2). The sequence of mitochondrial protein-coding genes is the same as that in other Lygaeoidea (Cao et al. 2020). Among the 37 genes, 23 genes (9 PCGs and 14 tRNAs) are on the majority strand (N-strand), while the remaining four PCGs, eight tRNAs, and two rRNA genes are on the minority strand (J-strand).

The basic composition of *P. porrectus* was A = 42.7%, T = 31.8%, G = 9.6%, and C = 15.8%, and of *P. sordidus*, A = 42.8%, T = 33.1%, G = 9.6%, C = 14.5%. Furthermore, both mitochondrial genome sequences were biased toward A and T. The AT content of *P. porrectus* was 63.74% and that of *P. sordidus* was 64.12%. The AT-skew value was greater than 0, whereas the GC skew value was less than 0, indicating that the base composition of *P. porrectus* and *P. sordidus* showed a strong A-bias and T-bias (Table 2).

Protein-coding genes

The complete length of the 13 PCGs of *P. porrectus* and *P. sordidus* were 10,991 bp and 10,993 bp, respectively. Of these, nine PCGs are located at the N-strand, and the other four PCGs were encoded on the J-strand (Fig. 2). Most PCGs started with ATN except for *cox1* and *nad4l* that began with TTG. Ten PCGs terminated with TAA/TAG, and the remaining three PCGs (*cox1*, *cox2*, and *cox3*)

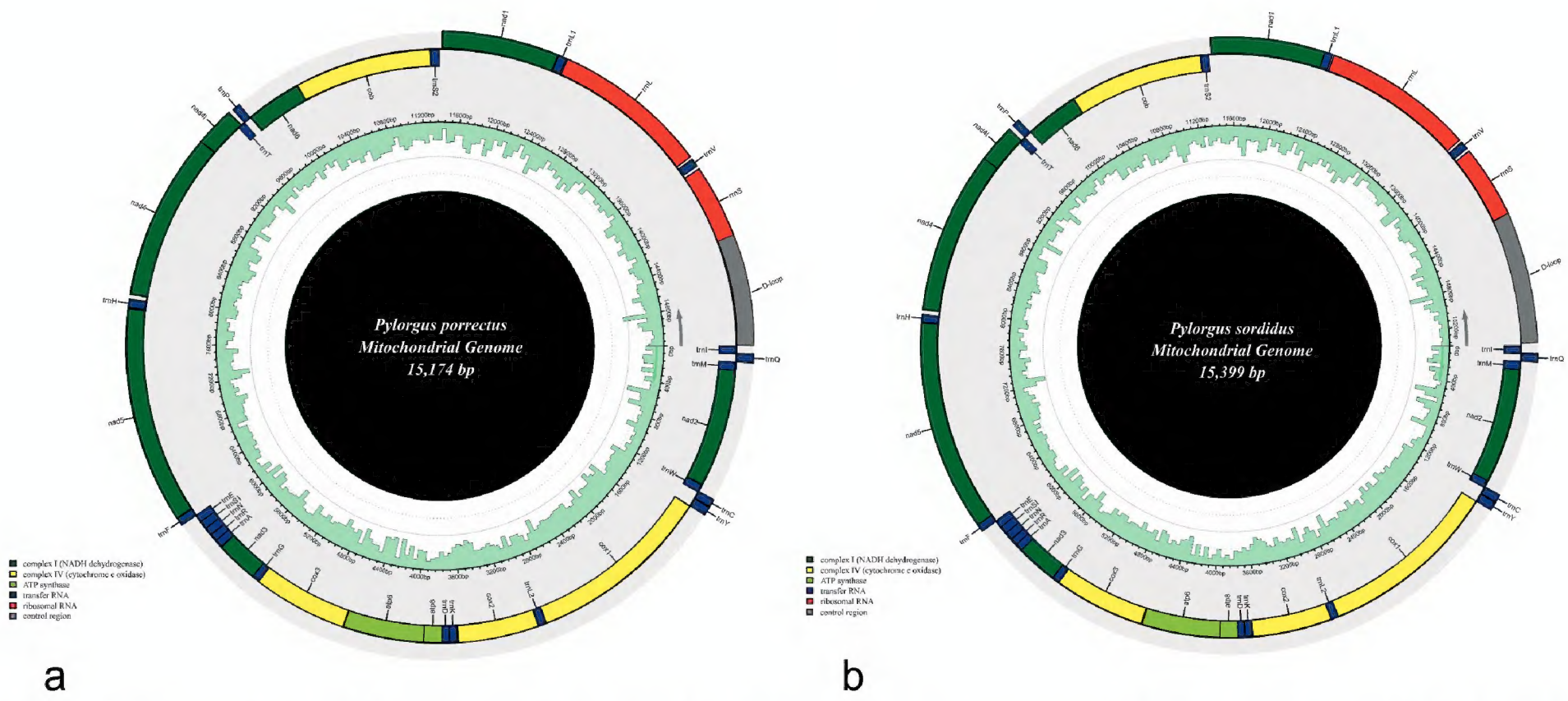

Figure 2. Circular maps of the complete mitogenome of *Pylorgus* species **a** *P. porrectus* **b** *P. sordidus*. Different colors indicate different types of genes and regions. Genes in the outer circle are located on the J-strand, and genes in the inner circle are located on the N-strand.

Table 2. Base content of the mitochondrial genome.

Gene	Size (bp)	A	T	G	C	A+T%	AT-skew	GC-skew
P. porrectus	15,174	42.7	31.8	9.6	15.8	74.5	0.15	−0.24
P. sordidus	15,399	42.8	33.1	9.6	14.5	75.9	0.13	−0.2

terminated with an incomplete T residue (Tables 3, 4). It has been speculated that these incomplete termination codons can be completed by adding 'A' during transcription (Ojala et al. 1981; Lavrov et al. 2002), and do not affect translation.

The RSCU of the two species was calculated (Fig. 3). The codons that were most used TTA-Leu and AGA-Arg. Most of the frequently used codons are composed of A and T, which may be related to the fact that the A-T skewness is higher than the G-C skewness in the PCGs of the two species.

Table 3. Mitochondrial composition of *Pylorgus porrectus*.

Gene	Position (bp)	Size (bp)	Strand	Direction	Intergenic nucleotides	Anti- or start/stop codons
trnI	1–64	64	N	Forward	0	
trnQ	62–130	69	J	Reverse	−3	
trnM	131–200	70	N	Forward	0	
nad2	201–1187	987	N	Forward	0	ATT/TAA
trnW	1178–1241	64	N	Forward	−10	
trnC	1234–1296	63	J	Reverse	−8	
trnY	1303–1370	68	J	Reverse	6	
cox1	1374–2907	1534	N	Forward	3	TTG/T–
trnL2	2908–2972	65	N	Forward	0	
cox2	2973–3648	676	N	Forward	0	ATA/T–
trnK	3649–3713	65	N	Forward	0	
trnD	3714–3774	61	N	Forward	0	
atp8	3775–3933	159	N	Forward	0	ATA/TAA
atp6	3927–4592	666	N	Forward	−7	ATG/TAA
cox3	4601–5378	778	N	Forward	8	ATT/T–
trnG	5379–5443	65	N	Forward	0	
nad3	5444–5797	354	N	Forward	0	ATA/TAG
trnA	5796–5860	65	N	Forward	−2	
trnR	5861–5920	60	N	Forward	0	
trnN	5923–5990	68	N	Forward	2	
trnS1	5990–6058	69	N	Forward	−1	
trnE	6058–6121	64	N	Forward	−1	
trnF	6122–6184	63	J	Reverse	0	
nad5	6185–7882	1698	J	Reverse	0	ATT/TAA
trnH	7886–7949	64	J	Reverse	3	
nad4	7987–9303	1317	J	Reverse	37	ATG/TAA
nad4l	9297–9605	309	J	Reverse	−7	TTG/TAA
trnT	9581–9643	63	N	Forward	−25	
trnP	9644–9706	63	J	Reverse	0	
nad6	9709–10170	462	N	Forward	2	ATC/TAA
cob	10170–11306	1137	N	Forward	−1	ATG/TAG
trnS2	11305–11374	70	N	Forward	−2	
nad1	11396–12319	924	J	Reverse	21	ATC/TAA
trnL1	12320–12384	65	J	Reverse	0	
rrnL	12392–13612	1221	J	Reverse	7	
trnV	13635–13700	66	J	Reverse	22	
rrnS	13726–14315	590	J	Reverse	25	

Table 4. Mitochondrial composition of *Pylorgus sordidus*.

Gene	Position (bp)	Size (bp)	Strand	Direction	Intergenic nucleotides	Anti- or start/stop codons
trnI	1–64	64	N	Forward	0	
trnQ	62–130	69	J	Reverse	−3	
trnM	131–201	71	N	Forward	0	
nad2	202–1188	987	N	Forward	0	ATT/TAA
trnW	1179–1242	64	N	Forward	−10	
trnC	1235–1297	63	J	Reverse	−8	
trnY	1305–1373	69	J	Reverse	10	
cox1	1377–2910	1534	N	Forward	3	TTG/T–
trnL2	2911–2975	65	N	Forward	0	
cox2	2976–3651	676	N	Forward	0	ATA/T–
trnK	3652–3716	65	N	Forward	0	
trnD	3717–3777	61	N	Forward	0	
atp8	3778–3936	159	N	Forward	0	ATA/TAA
atp6	3930–4595	666	N	Forward	−7	ATG/TAA
cox3	4604–5381	778	N	Forward	8	ATT/T–
trnG	5382–5444	63	N	Forward	0	
nad3	5445–5798	354	N	Forward	0	ATA/TAG
trnA	5797–5860	64	N	Forward	−2	
trnR	5861–5920	60	N	Forward	0	
trnN	5923–5990	68	N	Forward	2	
trnS1	5990–6058	69	N	Forward	−1	
trnE	6058–6121	64	N	Forward	−1	
trnF	6122–6184	63	J	Reverse	0	
nad5	6185–7882	1698	J	Reverse	0	ATT/TAA
trnH	7886–7949	64	J	Reverse	3	
nad4	7988–9304	1317	J	Reverse	38	ATG/TAA
nad4l	9298–9606	309	J	Reverse	−7	TTG/TAA
trnT	9582–9644	63	N	Forward	−25	
trnP	9645–9707	63	J	Reverse	0	
nad6	9710–10171	462	N	Forward	2	ATC/TAA
cob	10171–11307	1137	N	Forward	−1	ATG/TAG
trnS2	11306–11375	70	N	Forward	−2	
nad1	11397–12320	924	J	Reverse	21	ATC/TAA
trnL1	12321–12385	65	J	Reverse	0	
rrnL	12397–13613	1217	J	Reverse	11	
trnV	13636–13701	66	J	Reverse	22	
rrnS	13727–14316	590	J	Reverse	25	

The nucleotide diversity (Pi) of the two species based on 13 PCGs was computed (Fig. 4) and ranged from 0.05 to 0.11. Among the PCGs, *nad3* (0.11) had the highest Pi values, and *nad4l* (0.05) had the lowest Pi values, which implies that *nad4l* is the most conserved gene in *Pylorgus*.

The ratios of Ka/Ks for each gene of the 13 PCGs were also computed (Fig. 5). All Ka/Ks values were less than 1 and ranged from 0.01 to 0.13, indicating that

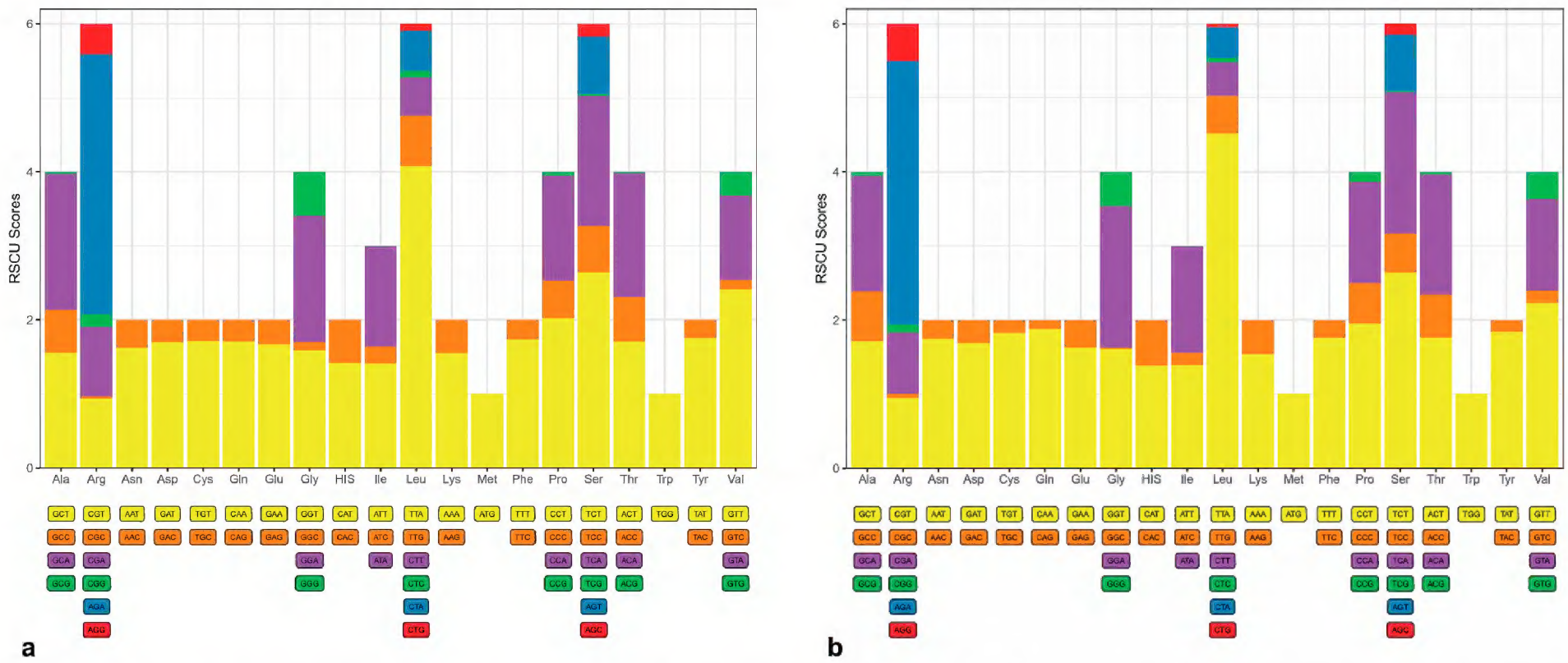

Figure 3. RSCU values of *Pylorgus* species **a** *P. porrectus* **b** *P. sordidus*. The abscissa represents the type of amino acid translated by the codon, and the ordinate represents the codon bias score calculated for the amino acid. The higher the score, the more the types of codons, and the more active the evolutionary variation of genes in the genome.

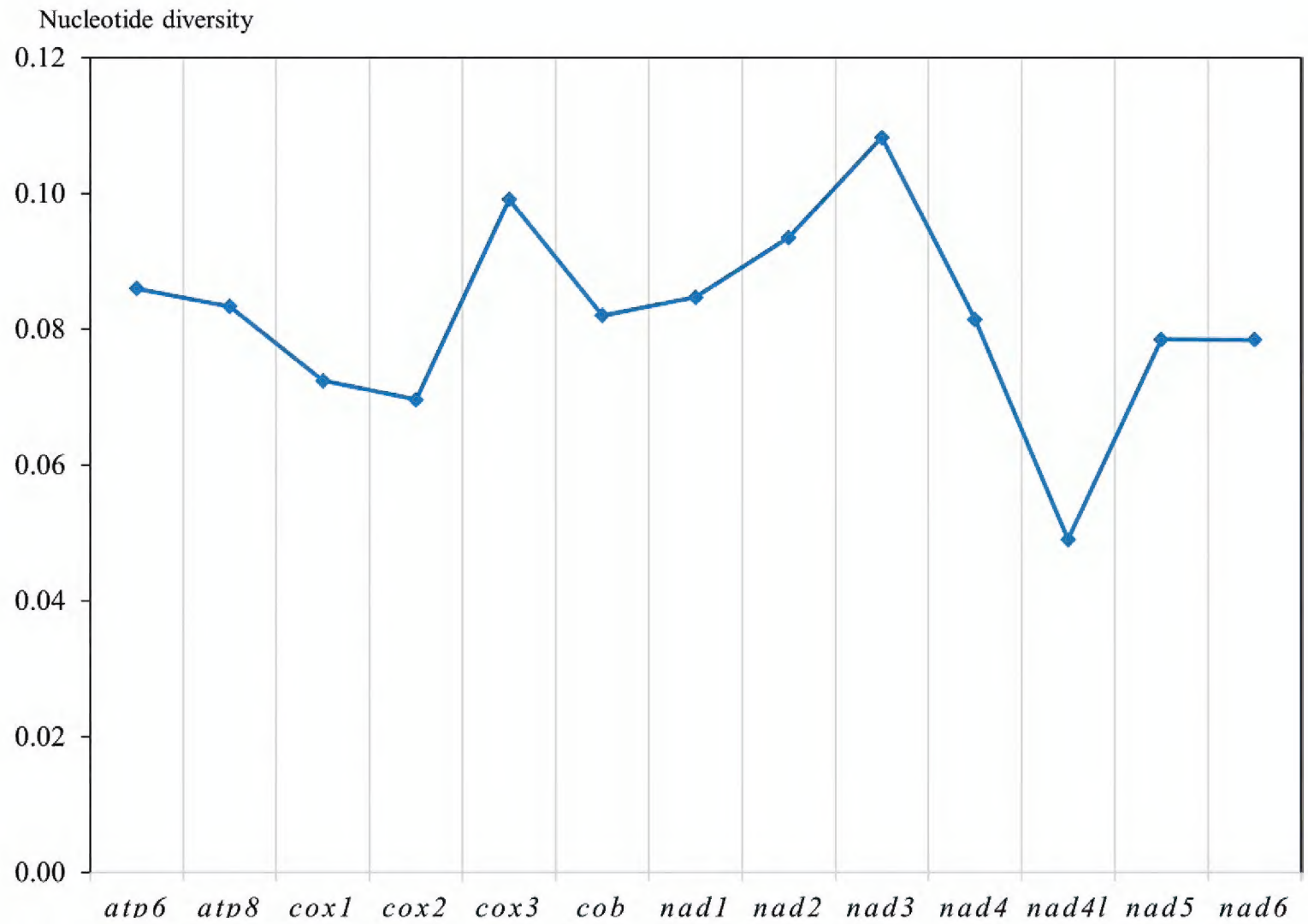

Figure 4. Nucleotide diversity (Pi) of 13 PCGs among two newly sequenced *Pylorgus* mitogenomes.

the genes have been subjected to purification selection. In particular, the Ka/Ks values were the highest for *nad4* and *nad5*, suggesting that they had the highest evolution speed, and lowest for *cox1*, indicating the slowest evolution.

Gene overlaps and intergenic spacers

Eleven gene overlaps were observed in the two mitogenomes, ranging from 1 bp to 25 bp (Tables 3, 4), and *nad4l* and *trnT* possessed the longest overlap.

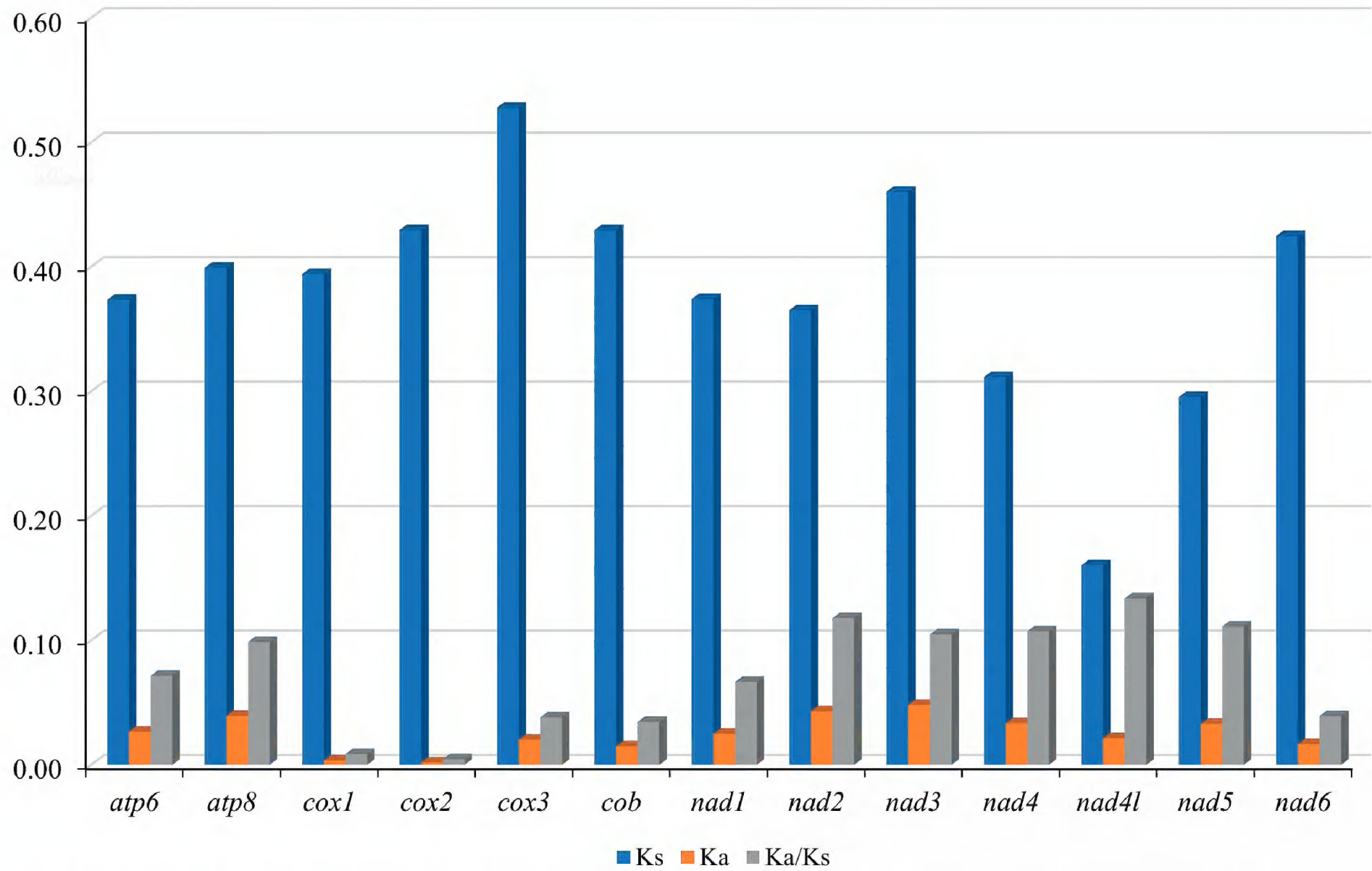

Figure 5. The ratios of Ka/Ks of 13 PCGs in the mitochondrial genomes of *Pylorgus porrectus* and *P. sordidus*.

Intergenic spacers were identified in the two mitogenomes, and their lengths ranged from 1 bp to 38 bp (Tables 3, 4). The longest intergenic spacer of 38 in *P. sordidus* was located between *trnH* and *nad4*.

Transfer RNA and ribosomal RNA genes

The two mitogenomes both contain the complete set of 22 tRNA genes typical of Lygaeidae mitogenomes, ranging from 60 to 71 bp, which is consistent with previously sequenced mitogenomes of Lygaeidae (Cao et al. 2020; Huang et al. 2021). Fourteen of the 22 tRNAs were on the N-strand, and eight were on the J-strand (Fig. 2).

All tRNA have the typical cloverleaf secondary structure, including the TΨC arm, the amino acid acceptor arm, the anticodon arm, and the dihydrouridine arm. Some of tRNA genes (*trnY*, *trnA*, *trnS1*, *trnF*, *trnH*, *trnP*, and *trnV*) showed individual base mismatches, which is a common phenomenon in insect mitogenomes (Zhang et al. 2019).

The *rrnL* genes of the two mitogenomes are located at the intergenic region between *trnL* and *trnV*, with lengths that range from 1217 bp to 1221 bp. The *rrnS* genes are located between *trnV* and the D-loop, which are both 590 bp in length. Both rRNAs are located on the N-strand.

Phylogenetic analysis

Phylogenetic relationships within Lygaeoidea were reconstructed based on mitochondrial 13 PCGs using BI and ML methods (Figs 6, 7). A total of 21 Lygae-

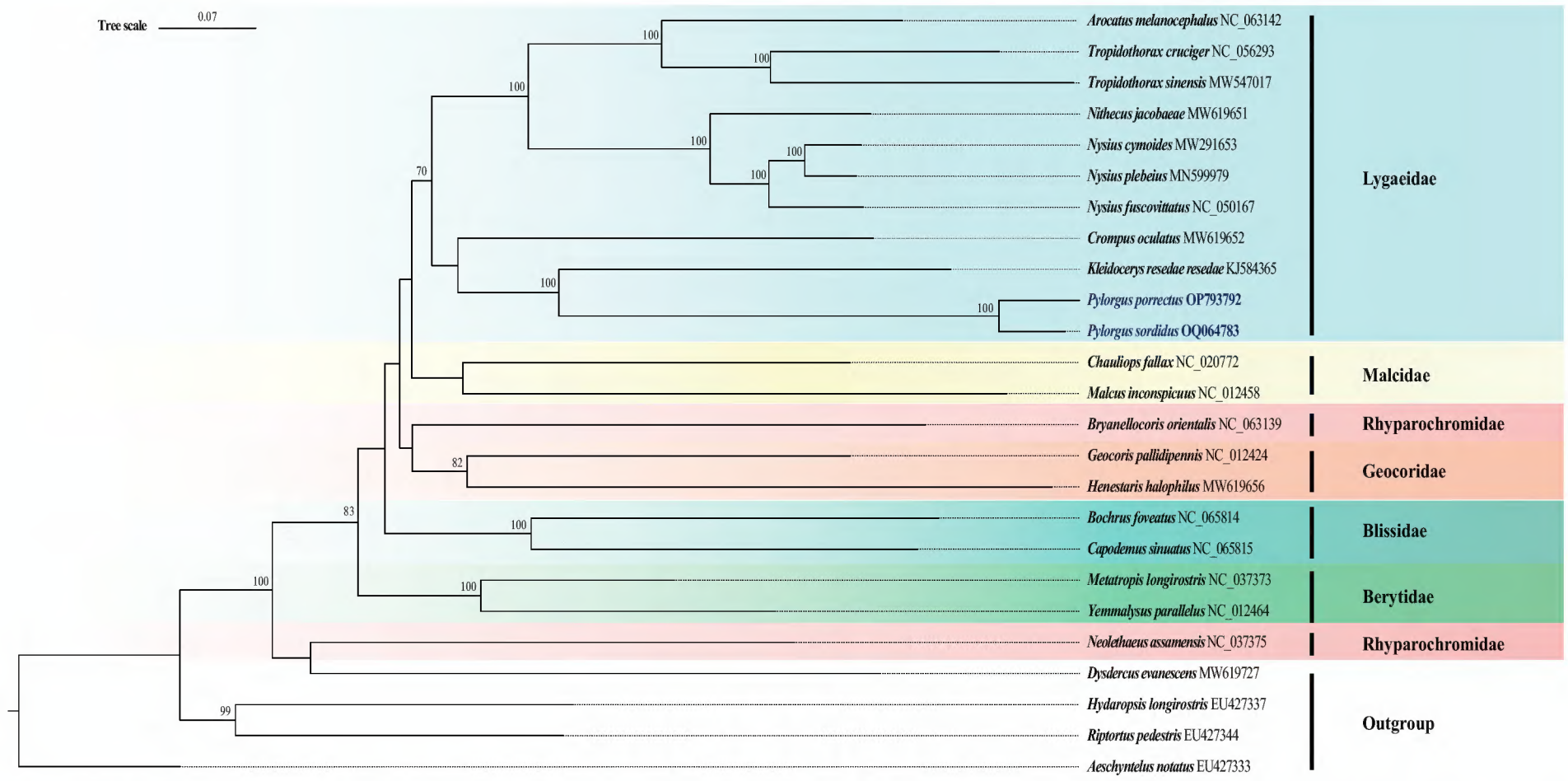

Figure 6. Phylogenetic tree inferred from ML methods based on 13 PCGs. Nodal support is given as standard bootstrap (%); only values > 70% are shown. The newly sequenced *Pylorgus porrectus* and *P. sordidus* mitogenomes are highlighted in dark blue and bold.

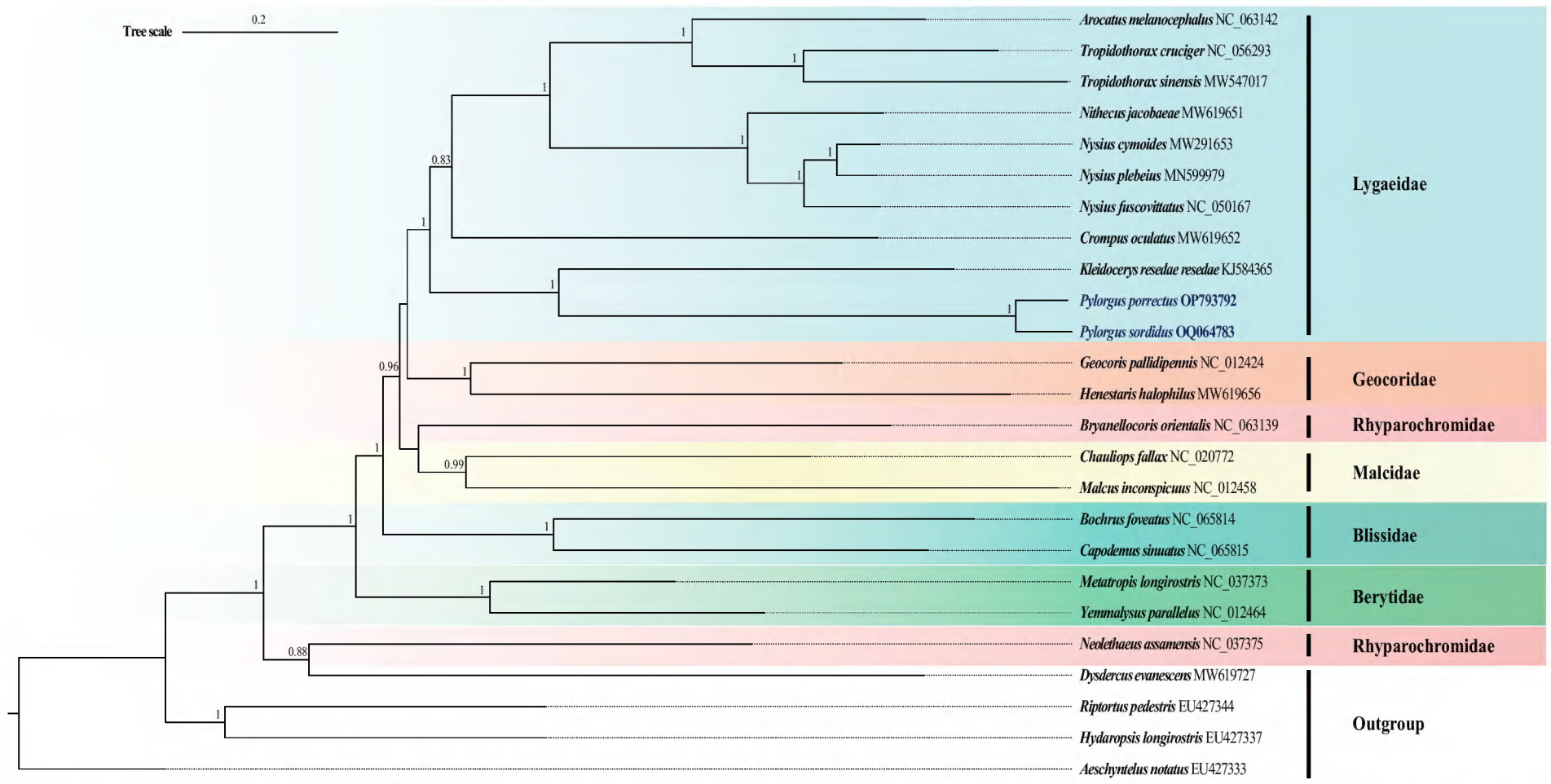

Figure 7. Phylogenetic tree inferred from BI methods based on 13 PCGs. Nodal support is given as partition model; only values > 0.80 are shown. The newly sequenced *Pylorgus porrectus* and *P. sordidus* mitogenomes are highlighted using dark blue and bold typeface.

oidea species were selected as the ingroup and an additional four species from Pyrrhocoroidea, Coreoidea, Rhopalidae, and Alydidae were used as the outgroup. Compared to the ML tree, the BI tree had higher confidence values, and the monophyly of all the studied families was supported except Rhyparochromidae.

The clades making up the Lygaeidae had high support values in the BI results and confirmed the monophyly of Lygaeidae (Figs 6, 7). The monophyly of Lygaeidae was also supported in the ML results, but the nodal support is not

so high. However, the Lygaeidae clusters as sister to Malcidae in the ML tree, but sister to Geocoridae in the BI tree, implying that the positions of Geocoridae and Malcidae are unstable. The two species of Rhyparochromidae are not clustered together. *Neolethaeus assamensis* clusters as sister to the Pyrrhocoroidea species, and together they are sister to the remaining ingroups.

Discussion

In this study, we sequenced and analyzed the mitogenomes of *Pylorgus porrectus* and *P. sordidus*, which had common and similar structures. The mitochondrial genome structure of the two *Pylorgus* species is a double-stranded closed loop, containing a non-coding control region sequence and encoding 37 genes. The two species showed a substantial nucleotide bias toward a higher A and T content, as do other Pentatomomorpha (Zhang et al. 2019; Cao et al. 2020; Huang et al. 2021; Carapelli et al. 2021; Xu et al. 2021; Zhu et al. 2023). All PCGs began with ATN except for *cox1* and *nad4l* that started with TTG. In total, 10 PCGs terminated with TAA/TAG and the remaining three PCGs (*cox1*, *cox2*, and *cox3*) terminated with incomplete T residues. The calculation of Ka/Ks values revealed that *nad4* and *nad5* had relatively higher evolutionary rates, and *cox1* was determined to be the most conserved gene. Eleven gene overlaps were observed in the two sequenced mitogenomes, and gene overlaps have also been found in other known Lygaeidae mitogenomes (Cao et al. 2020). All tRNA molecules have a typical cloverleaf structure (Li et al. 2017).

The phylogenetic results using 13 PCGs confirm the monophyly of Lygaeidae, which support the opinions of Henry (1997), Henry et al. (2015), and Schuh and Weirauch (2020). The ML tree shows that the topology within Lygaeidae is Ischnorhynchinae + (Lygaeinae + Orsillinae) (Fig. 6; Table 1). This result is in agreement with Cao et al. (2020) and Carapelli et al. (2021) but differs slightly from Henry's (1997) morphological hypothesis of Lygaeinae + (Ischnorhynchinae + Orsillinae). However, in the ML tree, *P. porrectus* and *P. sordidus* cluster with *Kleidocerys resedae* and then *Crompus oculatus* of Ischnorhynchinae (Fig. 6; Table 1), whereas in the BI tree, *P. porrectus* and *P. sordidus* only cluster with the *K. resedae*, and *C. oculatus* clusters with the other species of Lygaeinae and Orsillinae (Fig. 7; Table 1). We think this is mainly because the limited number of published mitogenomes within the Lygaeidae. This problem could be solved by sequencing additional mitogenomes of lygaeid species. The two selected species of Rhyparochromidae are not clustered together, which is similar with the results of Cao et al. (2020), Carapelli et al. (2021), and Huang et al. (2021). *Neolethaeus assamensis* clusters sister to the Pyrrhocoroidea species, and they together sister to the remaining ingroups in our result. More mitochondrial genomes need to be determined to better understand the monophyly of Rhyparochromidae. Overall, our results enrich the understanding of mitochondrial genome structure in the Lygaeidae and further supports the monophyly of the family containing the three subfamilies Ischnorhynchinae, Lygaeinae, and Orsillinae.

Acknowledgements

We are grateful to Shengchang Lai (Nanjing Forestry University) for helping with data analysis. We also thank Suyan Cao (Nanjing Forestry University) for

collecting specimens. We thank Thomas J. Henry (National Museum of Natural History, Washington DC.), the two anonymous reviewers, and the subject editor, Jader Oliveira, for their helpful and constructive comments. We also thank Guangyu Yu (Jiangxi Agricultural University) for revising the manuscript.

Additional information

Conflict of interest

No conflict of interest was declared.

Ethical statement

No ethical statement was reported.

Funding

This research was funded by Major Project of Agricultural Biological Breeding (grant no. 2022ZD0401501), the National Natural Science Foundation of China (grant no. 31402010), and the Highly Educated Talents Foundation in Nanjing Forestry University (grant no. G2014002).

Author contributions

Conceptualization, C.G. and W.D.; methodology, C.G. and W.D.; investigation, C.G. and W.D.; funding acquisition, C.G.; writing—original draft preparation, W.D.; writing—review and editing, C.G. Both authors have read and agreed to the published version of the manuscript.

Author ORCIDs

Cuiqing Gao https://orcid.org/0000-0002-0177-5161
Wen Dong https://orcid.org/0009-0004-6559-808X

Data availability

All of the data that support the findings of this study are available in the main text or Supplementary Information.

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A new genome assembly algorithm and its applications to Single-Cell Sequencing. Journal of Computational Biology 19(5): 455–477. https://doi.org/10.1089/cmb.2012.0021

Boetzer M, Pirovano W (2012) Toward almost closed genomes with GapFiller. Genome Biology 13(6): R56. https://doi.org/10.1186/gb-2012-13-6-r56

Cao Y, Wu HT, Li M, Chen WT, Yuan ML (2020) The complete mitochondrial genome of *Nysius fuscovittatus* (Hemiptera: Lygaeidae). Mitochondrial DNA, Part B, Resources 5(3): 3483–3484. https://doi.org/10.1080/23802359.2020.1827062

Carapelli A, Brunetti C, Cucini C, Cardaioli E, Soltani A, Amri M, Fanciulli PP, Nardi F (2021) The mitogenome of the true bug *Nysius cymoides* (Insecta, Heteroptera) and the phylogeny of Lygaeoidea. Mitochondrial DNA, Part B, Resources 6(8): 2366–2368. https://doi.org/10.1080/23802359.2021.1951139

Chen SF, Zhou YQ, Chen YR, Gu J (2018) Fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics (Oxford, England) 34(17): i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Dellapé PM, Henry TJ (2020) Lygaeoidea Species File. Version 5.0/5.0. http://Lygaeoidea.SpeciesFile.org [Accessed on 15 May 2023]

Eyles AC, Malipatil MB (2010) *Nysius caledoniae* Distant, 1920 (Hemiptera: Heteroptera: Orsillidae) a recent introduction into New Zealand, and keys to the species of *Nysius*, and genera of Orsillidae, in New Zealand. Zootaxa 2484(1): 45–52. https://doi.org/10.11646/zootaxa.2484.1.4

Ge XY, Li JL (2019) Review of the genus *Nysius* Dallas from Mongolian Plateau (Hemiptera: Heteroptera: Orsillidae). Zootaxa 4560(1): 171–183. https://doi.org/10.11646/zootaxa.4560.1.10

Henry TJ (1997) Phylogenetic analysis of family groups within the infraorder Pentatomomorpha (Hemiptera: Heteroptera), with emphasis on the Lygaeoidea. Annals of the Entomological Society of America 90(3): 275–301. https://doi.org/10.1093/aesa/90.3.275

Henry TJ (2000) Stilt bugs (Berytidae). In: Schaefer CW, Panizzi AR (Eds) Heteroptera of Economic Importance. CRC Press, Boca Raton, 725–735.

Henry TJ (2017) Biodiversity of Heteroptera. Wiley-Blackwell, Hoboken, 867 pp.

Henry TJ, Dellapé PM, de Paula AS (2015) The big-eyed bugs, chinch bugs, and seed bugs (Lygaeoidea). In: Panizzi AR, Grazia J (Eds) True Bugs (Heteroptera) of the Neotropics in Focus 2. Springer, Dordrecht, 459–514. https://doi.org/10.1007/978-94-017-9861-7_16

Huang WD, Gong SY, Wu YF, Song F, Li H (2021) The complete mitochondrial genome of *Tropidothorax sinensis* (Reuter, 1888) (Hemiptera: Lygaeidae). Mitochondrial DNA. Part B, Resources 6(7): 1808–1809. https://doi.org/10.1080/23802359.2021.1934156

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35(6): 1547–1549. https://doi.org/10.1093/molbev/msy096

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2016) PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular Biology and Evolution 34(3): 772–773. https://doi.org/10.1093/molbev/msw260

Lavrov DV, Boore JL, Brown WM (2002) Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: Duplication and nonrandom loss. Molecular Biology and Evolution 19(2): 163–169. https://doi.org/10.1093/oxfordjournals.molbev.a004068

Li H, Leavengood Jr JM, Chapman EG, Burkhardt D, Song F, Jiang P, Liu JP, Zhou XG, Cai W (2017) Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. Physical and Biological Sciences 284(1862): e20171223. https://doi.org/10.1098/rspb.2017.1223

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Malipatil MB (2010) Review and revision of *Nysius* Dallas of Australia and South-West Pacific (Hemiptera: Heteroptera: Orsillidae). Zootaxa 2410(1): 29–44. https://doi.org/10.11646/zootaxa.2410.1.2

Malipatil MB, Gao CQ, Eow LX (2020) Australian Lygaeoidea (Heteroptera) of Economic Importance Identification of Families, Tribes, and Representative Genera. Department of Jobs, Precincts and Regions, Melbourne, 170 pp.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32(1): 268–274. https://doi.org/10.1093/molbev/msu300

Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290(5806): 470–474. https://doi.org/10.1038/290470a0

Perna NT, Kocher TD (1995) Unequal base frequencies and the estimation of substitution rates. Molecular Biology and Evolution 12(2): 359–361. https://doi.org/10.1093/oxfordjournals.molbev.a040211

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61(3): 539–542. https://doi.org/10.1093/sysbio/sys029

Schuh RT, Slater JA (1995) True Bugs of the World (Hemiptera: Heteroptera): Classification and Natural History. Cornell University Press, Ithaca, 336 pp.

Schuh RT, Weirauch C (2020) True Bugs of the World (Hemiptera: Heteroptera): Classification and Natural History (2nd Edn.). Siri Scientific Press, Manchester, 800 pp.

Sweet MH (2000) Seed and chinch bugs (Lygaeoidea). In: Schaefer CW, Panizzi AR (Eds) Heteroptera of Economic Importance. CRC Press, Boca Raton, 143–264. https://doi.org/10.1201/9781420041859.ch6

Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology 56(4): 564–577. https://doi.org/10.1080/10635150701472164

Xu S, Wu Y, Liu Y, Zhao P, Chen Z, Song F, Li H, Cai W (2021) Comparative mitogenomics and phylogenetic analyses of Pentatomoidea (Hemiptera: Heteroptera). Genes 12(9): 1306. https://doi.org/10.3390/genes12091306

Yao Y, Ren D, Rider DA, Cai WZ (2012) Phylogeny of the infraorder Pentatomomorpha based on fossil and extant morphology, with description of a new fossil family from China. PLoS ONE 7(5): e37289. https://doi.org/10.1371/journal.pone.0037289

Zhang QL, Feng RQ, Li M, Guo ZL, Zhang LJ, Luo FZ, Cao Y, Yuan ML (2019) The complete mitogenome of *Pyrrhocoris tibialis* (Hemiptera: Pyrrhocoridae) and phylogenetic implications. Genes 10(10): 820. https://doi.org/10.3390/genes10100820

Zhang D, Gao FL, Jakovlic I, Zou H, Zhang J, Li WX, Wang GT (2020) PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources 20(1): 348–355. https://doi.org/10.1111/1755-0998.13096

Zhu WL, Yang L, Long JK, Chang ZM, Gong N, Mu YL, Lv SS, Chen XS (2023) Characterizing the complete mitochondrial genomes of three bugs (Hemiptera: Heteroptera) harming bamboo. Genes 14(2): 342. https://doi.org/10.3390/genes14020342